The dynamics of traded value revisited

Zoltán Eisler* and János Kertész†

Department of Theoretical Physics, Budapest University of Technology and Economics, Budapest, Hungary
(Dated: February 2, 2008)

We conclude from an analysis of high resolution NYSE data that the distribution of the traded value f_i (or volume) has a finite variance σ_i for the very large majority of stocks i, and the distribution itself is non-universal across stocks. The Hurst exponent of the same time series displays a crossover from weakly to strongly correlated behavior around the time scale of 1 day. The persistence in the strongly correlated regime increases with the average trading activity $\langle f_i \rangle$ as $H_i = H_0 + \gamma \log \langle f_i \rangle$, which is another sign of non-universal behavior. The existence of such liquidity dependent correlations is consistent with the empirical observation that $\sigma_i \propto \langle f_i \rangle^{\alpha}$, where α is a non-trivial, time scale dependent exponent.

The recent years have seen a number of important contributions of physics to various areas, among them to finance [1, 2]. The application of physical concepts often seems well suited to the analysis of financial time series, however, it is not without caveats. Often, the theoretical background of these methods is deeply rooted in physical laws that – naturally – do not apply to stock markets. In particular, observations regarding power laws [3], universality [4], and other empirical regularities [5] are often criticized. We carried out a thorough study of the traded value per unit time [6, 7, 8] and have arrived at the result that some earlier conclusions have to be modified. Here we present an analysis of some new data, which supports our earlier findings.

The paper is organized as follows. Section I introduces notations. Section II shows that the distribution of traded volume/value is not universal, and it is not in the Levy stable regime as suggested by Ref. [9]. Section III shows, that traded value displays only weak correlations for time scales shorter than one day. On longer horizons there is stronger persistence whose degree depends logarithmically on the liquidity of the stock. Finally, Section IV surveys the concept of fluctuation scaling, shows how it complements the observed liquidity dependence of correlations, and how those two form a consistent scaling theory.

I. NOTATIONS AND DATA

For a fixed time window size Δt, let us denote the total traded value of the ith stock at time t by

$$f_i^{\Delta t}(t) = \sum_{n, t_i(n) \in [t, t+\Delta t]} V_i(n), \qquad (1)$$

*Electronic address: eisler@maxwell.phy.bme.hu
†Also at Laboratory of Computational Engineering, Helsinki University of Technology, Espoo, Finland

where $t_i(n)$ is the time when the nth transaction of the ith stock takes place. Tick-by-tick data are denoted by $V_i(n)$, this is the value traded in transaction n, calculated as the product of the price and the traded volume.

Since price changes very little from trade to trade while variations of trading volume are much faster, the fluctuations of the traded value $f_i(t)$ are basically determined by those of traded volume. Price merely acts as a weighting factor that enables one to compare different stocks, while this also automatically corrects the data for stock splits and dividends. The correlation properties and the normalized distribution are nearly indistinguishable between traded volume and traded value.

This study is based on the complete Trades and Quotes database of New York Stock Exchange for the period $1994 - 1995$.

Note that throughout the paper we use 10-base logarithms.

II. TRADED VALUE DISTRIBUTIONS REVISITED

In this section, we first revisit the analysis done in Ref. [9]. That work finds that the cumulative distribution function of traded volume for time windows of $\Delta t = 15$ minutes decays as a power-law with a tail exponent $\lambda = 1.7 \pm 0.1$ for a wide range of stocks. This is the so called *inverse half cube law*, and it can be written as

$$\mathbb{P}_{\Delta t}(f) \propto f^{-(\lambda+1)}, \qquad (2)$$

where $\mathbb{P}_{\Delta t}$ is the probability density function of the same quantity.

The estimation of tail exponents is often difficult due to poor statistics of rare events, large stock-to-stock variations and the presence of correlations. For the same $1994 - 1995$ period of data and the same 15 minute time window certain stocks have λ values significantly higher than 1.7 [see Fig. 1(left)]. The tails of these distributions can be fitted by a power law over an order of magnitude, for the top $3 - 10\%$ of the events. The exponent

λ is around 2.8 for these examples. The question arises: Which value (if any) is correct?

In order to address this question we carried out a systematic investigation comprising the 1000 stocks with the highest total traded value in the TAQ database. We used variants of Hill's method [10, 11] to estimate the typical tail exponent, see Ref. [7] for details. The results of this Section are summarized in Table I. Note that in all cases the U-shaped intraday pattern of trading activity was removed.

Most descendants of Hill's method, including the ones applied here, contain a free parameter, namely the fraction p of top events to be considered to belong to the tail of the distribution (see Ref. [11] and refs. therein). According to Fig. 1(left) this should be set around $p \approx 3 - 10\%$.

First, let us follow the methodology of Ref. [9]. In that paper, the authors first they deduct the mean from the time series by taking $f_i(t) - \langle f_i \rangle$, where $\langle \cdot \rangle$ denotes time averaging. Then this series is used to estimate value of λ by applying Hill's method [12]. The choice $p = 0.03$ provides results in line with Ref. [9], for $\Delta t = 15$ min time windows one finds $\lambda = 1.67 \pm 0.20$. There are several issues with this approach:

1. p is a parameter that can be chosen arbitrarily. With the variation of p the same procedure can produce estimates ranging from $\lambda = 1.1 \pm 0.2$ ($p = 0.10$) to $\lambda = 2.15 \pm 0.2$ ($p = 0.005$).

2. The transformation significantly decreases the estimates of λ, down to the range of Levy stable distributions ($\lambda < 2$). Estimates for the untransformed data are given in Table I for comparison.

It is simple to show, that the first issue emerges, i.e. the estimates systematically depend on p, when one applies Hill's method to a finite sample from a distribution of the form

$$\mathbb{P}_{\Delta t}(f) \propto (f + f_0)^{-(\lambda+1)}, \qquad (3)$$

where f_0 is a non-zero constant. The transformation to $f_i(t) - \langle f_i \rangle$ does not resolve the problem, but biases the estimates further.

Instead, to correct for these biases one can (i) either find the proper constant f_0, remove it from the data, and apply Hill's estimator afterwards (ii) or apply the estimator of Fraga-Alves [11], which is insensitive to such shifts. Both of these estimates were found to be significantly higher [21]: $\lambda > 2$, see Table I. The methods are described in detail in Ref. [7].

The two corrected estimators show a strong tendency of increasing λ with increasing Δt. Monte Carlo simulations on surrogate datasets show that this is beyond what could be explained by decreasing sample size. For distributions with $\lambda < 2$ increasing window size should result in a convergence to the corresponding Levy distribution, and the measured λ's should be independent of Δt. Only when $\lambda > 2$ can the measured effective value

of λ systematically increase with Δt. For the 95% of the stocks the increasing tendency is observed and for a window size $\Delta t = 1$ day the respective λ's are greater than 2. These are strong indications that the distributions are not in the Levy stable regime, and thus the second moment exists.

Note that our calculations *assume* that the variable is asymptotically distributed as (3) and do not *prove* it. Still, the existence of the second moment is guaranteed by the absence of convergence to a Levy distribution. Consequently, it is possible to define the Hurst exponent for $f_i(t)$.

Regardless of the absence of the convergence to Levy stability there are qualitative similarities in the shape of the traded value distributions of various stocks [cf. Fig. 1(left)]. Nevertheless, the existence of a universal distribution can be rejected by a simple test [22].

If the form of the normalized distribution was universal, then the ratio of the standard deviation and the mean would have to obey $\sigma_i / \langle f_i \rangle = h$, where h is a constant independent of the stock. Equivalently, a relationship

$$\sigma_i \propto \langle f_i \rangle^{\alpha} \qquad (4)$$

would have to hold with an exponent $\alpha = 1$, at least on average. Even though one finds a monotonic dependence between the two quantities [as shown in Fig. 1(right)], the exponent is significantly less than 1. This means, that the ratio $\sigma / \langle f \rangle$ decreases with growing $\langle f \rangle$, i.e., the normalized distribution of f is narrower for larger stocks, so their trading exhibits smaller relative fluctuations. We will return to this observation in Section IV.

III. NON-UNIVERSALITY OF CORRELATIONS IN TRADED VALUE TIME SERIES

One of the classical tools of both financial analysis and physics is the measurement of the correlation properties of time series [1, 2, 13]. In particular, scaling methods [14] have a long tradition in the study of physical systems, where the Hurst exponent H_i is often calculated. For the traded value time series $f_i^{\Delta t}(t)$ of stock i this is defined as

$$\sigma_i^2(\Delta t) = \left\langle \left[f_i^{\Delta t}(t) - \left\langle f_i^{\Delta t}(t) \right\rangle \right]^2 \right\rangle \propto \Delta t^{2H_i}, \qquad (5)$$

Note that it follows from the results of Section II that the variance on the left hand side exists regardless of stock and for any window size Δt.

The measurements were carried out for all 2474 stocks that were continuously available on the market during $1994 - 1995$ [23]. Then we sorted the stocks into 6 groups according to the order of magnitude of their average traded value: $0 \le \langle f \rangle \le 10^4$, $10^4 \le \langle f \rangle \le 10^5$, ..., $10^8 \le \langle f \rangle$, all values in USD/min. Finally we averaged $\sigma_i^2(\Delta t)$ within each group. The obtained scaling plots are shown in Fig. 2.

Δt	Hill's λ ($p = 0.06$)	Ref. [9], $p = 0.03$	Shifted Hill's λ	$f_0/\langle f \rangle$	Fraga Alves ($p = 0.1$)
1 min	1.43 ± 0.09	1.45 ± 0.10	2.15 ± 0.15	3.0	1.98 ± 0.25
5 min	1.56 ± 0.13	1.55 ± 0.15	2.29 ± 0.25	2.8	2.04 ± 0.25
15 min	1.71 ± 0.20	1.67 ± 0.20	2.55 ± 0.35	2.8	2.1 ± 0.3
60 min	2.06 ± 0.30	1.90 ± 0.25	2.85 ± 0.45	1.8	2.1 ± 0.4
120 min	2.3 ± 0.4	2.0 ± 0.3	3.15 ± 0.70	1.6	2.1 ± 0.4
390 min	2.7 ± 0.6	2.1 ± 0.5	3.7 ± 0.9	1.2	no estimate

Table I: Median of the tail exponents of traded value calculated by four methods for $1994 - 1995$. The width of the distributions is given with the half distance of the 25% and 75% quantiles.
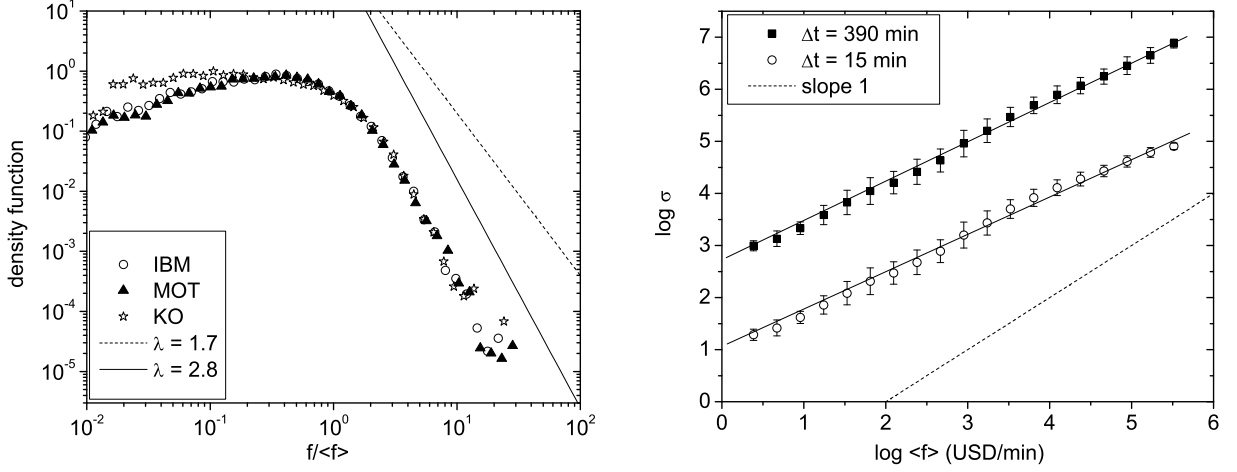

Figure 1: (left) The distribution of traded value in $\Delta t = 15$ min time windows, normalized by the mean. The plot displays three example stocks for the period $1994 - 1995$. The dashed and solid diagonal lines represent power-law distributions with tail exponents $\lambda = 1.7$ and 2.8, respectively. (right) The standard deviation σ_i of trading activity plotted versus the mean $\langle f_i \rangle$ of the same quantity. Both for $\Delta t = 15$ min, and $\Delta t = 390$ min ≈ 1 trading day, there is a power law relationship with slopes $\alpha(15 \text{ min}) = 0.74$, and $\alpha(390 \text{ min}) = 0.78$. A linear proportionality would support the existence of a universal distribution. However, the sublinear scaling relationship suggests that relative fluctuations in trading activity are smaller for stocks with higher average liquidity. *Note*: For better visibility, stocks were binned according to $\langle f \rangle$ and their $\log \sigma$ was averaged. The error bars correspond to the characteristic range within the bins.

All stocks display a crossover around window sizes of $\Delta t = 60 - 390$ min, and there are two sets of Hurst exponents: H_i^- valid below, and H_i^+ above the crossover. These characterize the strength of intraday and long time correlations, respectively. The behavior on these two time scales is very different.

1. For intraday fluctuations, regardless of stock $H^- \approx 0.51 - 0.52$. This means that intraday fluctuations of traded value are nearly uncorrelated.

2. For long time fluctuations the data are correlated, but the strength of correlations depends strongly on the liquidity of the stock. As one moves to groups of larger $\langle f \rangle$, the strength of correlations (H^+) grows, up to $H^+ \approx 0.8$.

3. If one shuffles the time series, correlations are destroyed, and $H_{\text{shuff}} = 0.5$.

The same phenomenon can be characterized by directly plotting the dependence of $H^\pm$ on $\langle f \rangle$, as done in Fig. 3. Such a dependence is well described by a logarithmic law:

$$H_i^\pm = H_0^\pm + \gamma^\pm \log \langle f_i \rangle, \qquad (6)$$

where $\gamma^- = 0.00 \pm 0.01$, and $\gamma^+ = 0.053 \pm 0.01$. For the shuffled time series $\gamma_{\text{shuff}} = 0$.

These results indicate, at least in the case of traded value, the absence of universal behavior. Liquidity (or, analogously, company size) is a relevant quantity, which acts as a *continuous* parameter of empirical observables, in particular the strength of correlations and the distribution of f. Related results can be found in Refs. [7, 15, 16, 17].
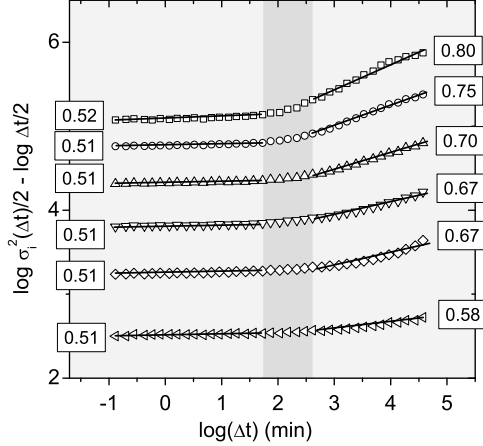
Figure 2: The normalized variance $\frac{1}{2}\log\sigma_i^2(\Delta t) - \frac{1}{2}\log\Delta t$ for the six groups of companies, with average traded values $\langle f\rangle \in [0, 10^4)$, $\langle f\rangle \in [10^4, 10^5)$, ..., $\langle f\rangle \in [10^8, \dots)$ USD/min, increasing from bottom to top. A horizontal line would mean the absence of autocorrelations in the data. Instead, one observes a crossover phenomenon in the regime $\Delta t = 60 - 390$ mins, indicated by darker background. Below the crossover all stocks show very weakly correlated behavior, $H^- \approx 0.5$. Above the crossover, the strength of correlations, and thus the slope corresponding to $H^+ - \frac{1}{2}$, increases with the liquidity of the stock. The asymptotic values of $H^\pm$ are indicated in the plot.

IV.   FLUCTUATION SCALING

Fluctuation scaling is a general phenomenon, observed in a wide range of complex systems [6, 8, 18]: A scaling law connects the standard deviation σ_i and the average $\langle f_i\rangle$ of the same quantity. In the case of the trading activity of stocks we have already presented this result in Section II [cf. Eq. (4)]. Now we give a more detailed discussion.

Let us start from our observation that

$$\sigma_i(\Delta t) \propto \langle f_i\rangle^{\alpha(\Delta t)}, \tag{7}$$

where the scaling variable is $\langle f_i\rangle$, or more appropriately the stock i, and Δt is kept constant [see Fig. 1(right)]. Notice that $\sigma_i(\Delta t)$ is the same as in the definition of the Hurst exponent in Eq. (5), where i was constant and Δt was varied.

In Eq. (7) the window size Δt is a free parameter. This scaling law persists for any Δt, but α strongly depends on its value, as shown in Fig. 4. For small time windows (up to 60 min), $\alpha(\Delta t) \approx 0.74$, then, after a crossover regime, when $\Delta t > 390$ min, there is a logarithmic trend. This can be summarized as

$$\alpha(\Delta t) = \alpha_0^\pm + \gamma^\pm \log\Delta t, \tag{8}$$

where $\cdot^\pm$ refers to the regimes $\Delta t < 60$ min and $\Delta t > 390$ min. The constants are $\alpha_0^- = 0.74$, $\gamma^- = 0$, and
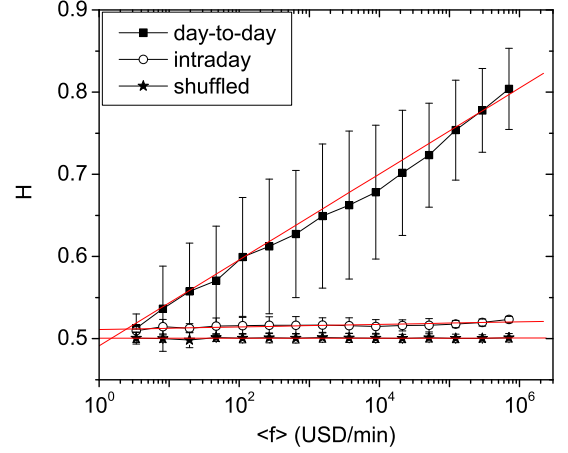

Figure 3: Value of the Hurst exponents of traded value for the time period $1994-1995$. For short time windows (O, $\Delta t < 60$ min) all signals are nearly uncorrelated, $H^- \approx 0.51 - 0.53$. The fitted slope is $\gamma^- = 0.00 \pm 0.01$. For larger time windows ($\blacksquare$, $\Delta t > 390$ min) the strength of correlations depends logarithmically on the mean trading activity of the stock, $\gamma^+ = 0.053 \pm 0.01$ for $1994 - 1995$. Shuffled data ($\bigstar$) display no correlations, thus $H_{\text{shuff}} = 0.5$, which also implies $\gamma_{\text{shuff}} = 0$. *Note*: Groups of stocks were binned, and their Hurst exponents were averaged. The error bars correspond to the standard deviations in the bins.

$\gamma^+ = 0.052 \pm 0.01$. For shuffled time series, $\alpha(\Delta t) = 0.74$ regardless of Δt, i.e., $\gamma_{\text{shuff}} = 0$.

A visual comparison of Figs. 2 and 4 reveals that the crossover in the behavior of $\alpha(\Delta t)$ and H falls into the same interval. Moreover, when $\Delta t < 60$ min, both $\alpha(\Delta t)$ and $H^-(\langle f\rangle)$ are constant. For $\Delta t > 390$ min, both $\alpha(\Delta t)$ and $H^+(\langle f\rangle)$ vary logarithmically with their arguments (see Figs. 3 and 4).

In order to better understand the connection between temporal correlations and fluctuation scaling, let us repeat here Eqs. (6) and (8):

$$\alpha(\Delta t) = \alpha_0^\pm + \gamma^\pm \log\Delta t,$$
$$H_i = H_0^\pm + \gamma^\pm \log\langle f_i\rangle.$$

Beyond the obvious symmetry of these two logarithmic laws, notice that the prefactors are equal: in both equations $\gamma^- \approx 0$ and $\gamma^+ \approx 0.05$.

It is easy to show [8] that none of this is a simple coincidence. If both fluctuation scaling and long range autocorrelations are present in data, there are only two possible ways for their coexistence:

1. Correlations are homogeneous throughout the system, $H_i = H_0$, $\gamma = 0$, and α is independent of Δt. This is realized for $\Delta < 60$ min. For shuffled time series correlations are absent, thus such data also fall into this category.
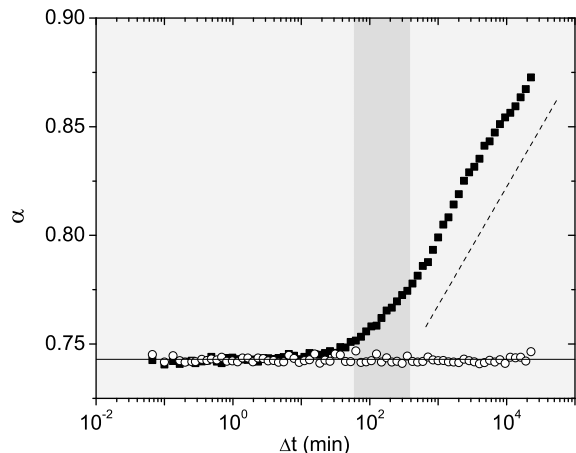
Figure 4: The dependence of the scaling exponent α on the window size Δt for the years $1994 - 1995$. The lighter shaded intervals have well-defined Hurst exponents and values of γ, the crossover is indicated with a darker background. Without shuffling (■) there are two linear regimes: For shorter windows $\alpha = 0.74 \pm 0.02$, the slope is $\gamma^- = \gamma(\Delta t < 60$ min$) = 0.00 \pm 0.01$ (solid line), while for longer windows α grows logaritmically, with a slope $\gamma^+ = \gamma(\Delta t > 390$ min$) = 0.052 \pm 0.01$ (dashed line). For shuffled data (O) the exponent is independent of window size, $\alpha(\Delta t) = 0.74 \pm 0.02$.

2. Both the $H(\langle f_i \rangle)$ and $\alpha(\Delta t)$ are logarithmic functions of their arguments with the same coefficient γ^+. This is realized for $\Delta > 390$ min.

In other words the coexistence of the two scaling laws is so restrictive, that if the strength of correlations depends on $\langle f \rangle$ at all, then the realized logarithmic dependence is the only possible scenario.

V. CONCLUSIONS

In this paper, we analyzed the empirical properties of trading activity on the New York Stock Exchange. We showed that, in contrast to earlier findings, the distribution of traded value is not in the Levy stable regime, and is not universal. Traded value is nearly uncorrelated on an intraday time scale, while on daily or longer scales fluctuations show strong persistence, whose strength grows logarithmically with the liquidity of the stock. This effect is in harmony with findings on fluctuation scaling, a general scaling framework for complex systems.

All our results imply, that the notion of universality must be used with extreme care in the context of financial markets, where the concepts and the theoretical background are radically different from those in physics. The liquidity of a stock strongly affects the distribution and the correlation structure of its trading activity. This dependence is continuous, which means the absence of universality classes in trading dynamics. The dynamical process responsible for such a dependence is yet to be identified.

Finally, we would like to make two remarks. Firstly, in Refs. [9] and [19] it is stated that the so called inverse half cubic law is observable not only on the 15 minute level but also in the tick-by-tick data. Our analysis dealt with data aggregated for 1 minute and more, and we showed that the assumption of a power law decay is not consistent with the inverse half cubic law for these cases. Secondly, in Ref. [19] a footnote mentions the possibility of an exponential cutoff in the distribution. This assumption would influence the estimators strongly and we did not consider this case. We thank the anonymous referee for calling our attention to these points.

[1] J.-P. Bouchaud and M. Potters, *Theory of Financial Risk* (Cambridge University Press, Cambridge, 2000).
[2] R.N. Mantegna and H.E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, 1999).
[3] M. Mitzenmacher, Internet Math **2**, 525 (2005).
[4] P. Ball, Nature **441**, 686 (2006).
[5] M. Gallegati, S. Keen, T. Lux, and P. Ormerod, Physica A **370**, 1 (2006).
[6] Z. Eisler, J. Kertész, S.-H. Yook, and A.-L. Barabási, Europhys. Lett. **69**, 664 (2005).
[7] Z. Eisler and J. Kertész, Eur. Phys. J. B **51**, 145 (2006),
(also arXiv:physics/0508156).
[8] Z. Eisler and J. Kertész, Phys. Rev. E **73** (2006).
[9] P. Gopikrishnan, V. Plerou, X. Gabaix, and H. Stanley, Phys. Rev. E **62**, 4493 (2000).
[10] B. Hill, Annals of Statistics **3**, 1163 (1975).
[11] M.I. Fraga Alves, Extremes **4**, 199 (2001).
[12] P. Gopikrishnan and V. Plerou, personal communication.
[13] M. Tumminello, T. Aste, T. di Matteo, and R. Mantegna, Proc. Natl. Acad. Sci. USA **102**, 10421 (2005).
[14] J. Kantelhardt, S. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, Physica A **316**, 87 (2002).

[15] Z. Eisler and J. Kertész, Europhys. Lett. **77**, 28001 (2007).

[16] G. Bonanno, F. Lillo, and R. Mantegna, Physica A **280**, 136 (2000).

[17] P. Ivanov, A. Yuen, B. Podobnik, and Y. Lee, Phys. Rev. E **69**, 56107 (2004).

[18] M. de Menezes and A.-L. Barabási, Phys. Rev. Lett. **92**, 28701 (2004).

[19] X. Gabaix, P. Gopikrishnan, V. Plerou, and H. Stanley, Quart. J. Econ. **121**, 461 (2006).

[20] F. Lillo and R. Mantegna, Phys. Rev. E **62**, 6126 (2000).

[21] The Fraga-Alves estimator converges very slowly, and it underestimates the actual values of λ from small samples. Its estimates can be interpreted as lower bounds for λ.

[22] Similar techniques were used in Refs. [7, 20] to show non-universality in the distribution of returns and intertrade times.

[23] For a similar analysis of the years $2000 - 2002$, see Ref. [7].